UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/626,856	07/25/2003	Thorsten H. Brants	A3052-US-NP XERZ 2 01563	9802

61962   7590   12/28/2007

FAY SHARPE / XEROX - PARC
1100 SUPERIOR AVENUE
SUITE 700
CLEVELAND, OH 44114

EXAMINER
LOVEL, KIMBERLY M

ART UNIT	PAPER NUMBER
2167	

MAIL DATE	DELIVERY MODE
12/28/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

		Application No.	Applicant(s)
		10/626,856	BRANTS ET AL.
Office Action Summary		Examiner	Art Unit	
		Kimberly Lovel	2167	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *20 September 2007*.

2a)☐ This action is **FINAL**.   2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-15 and 35-38* is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-15 and 35-38* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☐ All   b)☐ Some * c)☐ None of:

        1.☐ Certified copies of the priority documents have been received.

        2.☐ Certified copies of the priority documents have been received in Application No. _____.

        3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1)☒ Notice of References Cited (PTO-892)

2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3)☐ Information Disclosure Statement(s) (PTO/SB/08)
    Paper No(s)/Mail Date _____.

4)☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____.

5)☐ Notice of Informal Patent Application

6)☐ Other: _____.

DETAILED ACTION

1.   Claims 1-15 and 35-38 are rejected. Claims 16-34 have been cancelled.

Continued Examination Under 37 CFR 1.114

2.   A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 20 September 2007 has been entered.

Claim Objections

3.   **Claim 38** is objected to because line 3 recites the phrase "one indirect **storing** characteristic". It seems that "storing" should be replaced with "story."

Appropriate correction is required.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of

the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of

the various claims was commonly owned at the time any inventions covered therein

were made absent any evidence to the contrary. Applicant is advised of the obligation

under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was

not commonly owned at the time a later invention was made in order for the examiner to

consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g)

prior art under 35 U.S.C. 103(a).

4.   **Claims 1, 7-15, 37 and 38 are rejected under 35 U.S.C. 103(a) as being**

unpatentable over the article "Topic Detection and Tracking Pilot Study Final

Report" by Allan et al (hereafter Final Report) in view of the article "Relevance

Models for Topic Detection and Tracking" by Lavrenko et al in view of US Patent

No 7,085,755 to Bluhm et al (hereafter Bluhm).

Referring to claim 1, Final Report discloses a computer-implemented method of detecting new events (see section 1.1: Background, lines 1-6) comprising the steps of:

determining at least one story characteristic based on at least one of:

an average story similarity story (see section 3.1: Detection Evaluation: lines 34-44) and a same event-same source story characteristic;

determining a story corpus, each story associated with at least one event (see section 1.2: The Corpus; and section 2.1: Evaluation, lines 11-16 – the corpus is the TDT study corpus (a source) has been developed by the TDT study effort and comprises of 16,000 stories from Reuters newswire and CNN broadcast news transcripts);

determining a new story associated with at least one event (see section 3: New Event Detection, lines 19-24 – a Yes/No decision per story made at the time when the story arrives regarding whether or not the story is the first reference to a newly reported event);

determining story-pairs based on the new-story and each story in the story corpus (see section 3.1: Detection Evaluation, lines 6-48 – each new story is associated with a cluster which is considered to represent a story-pair);

determining at least one inter-story similarity metric [score] for the story-pairs (see section 3.1: Detection Evaluation, lines 30-33 – the score indicates how confident the system is that the story being processed discusses the cluster event); and

outputting a new story event indicator [YES or NO decision] (see section 3.1: Detection Evaluation, lines 23-29) if the event associated with the new story is similar to

the events associated with the source-identified story corpus based on the inter-story

similarity metrics (see section 3.1: Detection Evaluation, lines 30-33 - score).

However, Final Report fails to explicitly teach the limitation of determining at least

one adjustment to the inter-story similarity metrics based on at least one story

characteristic and then using the adjustment to output a new story indicator. Relevance

Model also discloses a method for detecting new events similar to that of Final Report

including the further limitation. In particular, Relevance Model discloses determining at

least one adjustment to the inter-story similarity metrics based on at least one story

characteristic (see section 4.3: Relevance Model Performance – the KL metric is

adjusted by a clarity value) and using the adjustment to output a new event story

indicator (see section 3.2: Measuring Topic Similarity).

It would have been obvious to one of ordinary skill at the time the invention was

made to utilize Relevance Model's method of determining at least one adjustment to the

inter-story similarity metrics based on at least one story characteristic with Final

Report's method of detecting new events by using the KL metric and the score metric.

One would have been motivated to do so since the Kullback-Leibler metric provides

relevance modeling, which enhances the topic model estimate associated with a news

story (Relevance Model: see abstract, lines 3-6).

While the combination of Final Report and Relevance Model (hereafter

FinalReport/Relevance Model) discloses a story corpus and a new story,

FinalReport/RelevanceModel fails to explicitly disclose the further limitations of a

source-identified story corpus and source-identified new story. Bluhm discloses

managing a large corpus of documents (see abstract), including the further limitations of a source-identified story corpus and a source-identified new story [source/author/publisher] (see column 6, lines 33-47; column 7, lines 27-42; and column 22, lines 1-10).

It would have been obvious to one of ordinary skill in the art at the time of the invention to utilize the identification of the source of the corpus and new story as disclosed by Bluhm with the corpus and new story of FinalReport/Relevance Model. One would have been motivated to do so in order to increase the accuracy of the results by decreasing the size of the document set searched by limiting the search to a particular source.

Referring to claim 7, the combination of FinalReport/RelevanceModel and Bluhm (hereafter FinalReport/RelevanceModel/Bluhm) teaches the method of claim 1, wherein the inter-story similarity metrics are comprised of: at least one story frequency model (Final Report: see section 4.1: Detection Experiments, lines 13-15); and

at least one event frequency model combined using terms weights (Relevance/Model: see section 5.1: Tracking algorithm, lines 1-23).

Referring to claim 8, FinalReport/RelevanceModel/Bluhm teaches the method of claim 1, wherein the inter-story similarity metrics are comprised of at least one story frequency model (Final Report: see section 4.1: Detection Experiments, lines 13-15); and

at least one story characteristic frequency model combined using terms weights (Relevance Model: see Figure 6).

Referring to claim 9, FinalReport/RelevanceModel/Bluhm teaches the method of claim 8, where the adjustments based on the story characteristics are applied to the term weights (Relevance Model: see section 1: Introduction, lines 29-31 and Figure 6).

Referring to claim 10, FinalReport/RelevanceModel/Bluhm teaches the method of claim 8, where the adjustments based on the story characteristics are applied to the inter-story similarity metrics (Relevance Model: see section 4.3: Relevance Model Performance – the KL metric is adjusted by a clarity value).

Referring to claim 11, FinalReport/RelevanceModel/Bluhm teaches the method of claim 1, wherein the inter-story similarity metrics are comprised of at least one term frequency-inverse event frequency model (Final Report: see section 4.1: Event Detection, lines 11-23) and where the events are classified based on at least one of: story labels and a predictive model (Relevance Model: see section 5.1: Tracking Algorithm, lines 11-23).

Referring to claim 12, FinalReport/RelevanceModel/Bluhm teaches the method of claim 8, wherein an event frequency is determined based on term t and ROI category $rmax$ from the formula: $ef_{r\max}(t) = \max_{r \in R}(ef(r,t))$ (Final Report: see section 3.2: Measuring Topic Similarity – the equation finds the probability of the topic).

Referring to claim 13, FinalReport/RelevanceModel/Bluhm teaches the method of claim 8, wherein an the inverse event frequency is determined based on term t, and events e and $rmax$ in the set of ROI categories from the formula: $IEF(t) = \log\left[\dfrac{N_{e,r}}{ef(r,t)}\right]$

(Final Report: see section 3.5 Results, Analysis, and Future Work – finding the inverse

document frequency is comparable to finding the inverse event frequency).

Referring to claim 14, FinalReport/RelevanceModel/Bluhm teaches the method

of claim 8, wherein an inverse event frequency is determined based on term t,

categories e, r and rmax in the set of ROI categories and P(r), the probability of ROI r

from the formula: $IEF'(t) = \sum_{r \in R} P(r) \log\left[\dfrac{N_{e,r}}{ef(r,t)}\right]$ (Final Report: see section 3.5 Results,

Analysis, and Future Work – finding the inverse document frequency is comparable to

finding the inverse event frequency; the derivative of the first equation has been taken).

Referring to claim 15, FinalReport/RelevanceModel/Bluhm teaches the method

of claim 1 further comprising the step of determining a subset of stories from the source-

identified story corpus and the source-identified new story based on at least one story

characteristic (Final Report: see section 3.1: Detection Evaluation – the stories within

the corpus are placed in clusters which is considered to represent a subset).

Referring to claim 37, FinalReport/RelevanceModel/Bluhm discloses the

computer-implemented method of claim 1, in which the new event indicator is displayed

on at least one of a visual, audio or tactile output device (Final Report: Section 3.6:

Open Issues).

Referring to claim 38, Final Report discloses a computer-implemented method of detecting new events (see section 1.1: Background, lines 1-6) comprising the steps of:

determining at least one direct story characteristic or one indirect storing characteristic based on at least one of: a story authorship, a story language, an average story similarity story (see section 3.1: Detection Evaluation: lines 34-44) and a same event-same source story characteristic;

determining a story corpus, each story associated with at least one event (see section 1.2: The Corpus; and section 2.1: Evaluation, lines 11-16 – the corpus is the TDT study corpus (a source) has been developed by the TDT study effort and comprises of 16,000 stories from Reuters newswire and CNN broadcast news transcripts);

determining a subset the story corpus (Final Report: see section 3.1: Detection Evaluation – the stories within the corpus are placed in clusters which is considered to represent a subset);

determining a new story associated with at least one event (see section 3: New Event Detection, lines 19-24 – a Yes/No decision per story made at the time when the story arrives regarding whether or not the story is the first reference to a newly reported event);

determining story-pairs based on the new-story and each story in the story corpus (see section 3.1: Detection Evaluation, lines 6-48 – each new story is associated with a cluster which is considered to represent a story-pair);

determining at least one inter-story similarity metric [score] for the story-pairs (see section 3.1: Detection Evaluation, lines 30-33 – the score indicates how confident the system is that the story being processed discusses the cluster event), wherein the inter-story similarity metrics are comprised of at least one story frequency model (Final Report: see section 4.1: Detection Experiments, lines 13-15); and

outputting a new story event indicator [YES or NO decision] (see section 3.1: Detection Evaluation, lines 23-29) if the event associated with the new story is similar to the events associated with the source-identified story corpus based on the inter-story similarity metrics (see section 3.1: Detection Evaluation, lines 30-33 - score).

However, Final Report fails to explicitly teach the limitations of at least one event frequency model combined using terms weights and determining at least one adjustment to the inter-story similarity metrics based on at least one story characteristic and then using the adjustment to output a new story indicator. Relevance Model also discloses a method for detecting new events similar to that of Final Report including the further limitation. In particular, Relevance Model discloses the further limitations of at least one event frequency model combined using terms weights (Relevance/Model: see section 5.1: Tracking algorithm, lines 1-23) and determining at least one adjustment to the inter-story similarity metrics based on at least one story characteristic (see section 4.3: Relevance Model Performance – the KL metric is adjusted by a clarity value) and using the adjustment to output a new event story indicator (see section 3.2: Measuring Topic Similarity).

It would have been obvious to one of ordinary skill at the time the invention was made to utilize Relevance Model's combined frequency model and method of determining at least one adjustment to the inter-story similarity metrics based on at least one story characteristic with Final Report's method of detecting new events by using the KL metric and the score metric. One would have been motivated to do so since the Kullback-Leibler metric provides relevance modeling, which enhances the topic model estimate associated with a news story (Relevance Model: see abstract, lines 3-6).

While the combination of Final Report and Relevance Model (hereafter FinalReport/Relevance Model) discloses a story corpus and a new story, FinalReport/RelevanceModel fails to explicitly disclose the further limitations of a source-identified story corpus and source-identified new story. Bluhm discloses managing a large corpus of documents (see abstract), including the further limitations of a source-identified story corpus and a source-identified new story [source/author/publisher] (see column 6, lines 33-47; column 7, lines 27-42; and column 22, lines 1-10).

It would have been obvious to one of ordinary skill in the art at the time of the invention to utilize the identification of the source of the corpus and new story as disclosed by Bluhm with the corpus and new story of FinalReport/Relevance Model. One would have been motivated to do so in order to increase the accuracy of the results by decreasing the size of the document set searched by limiting the search to a particular source.

5.   Claim 2 is rejected under 35 U.S.C. 103(a) as being unpatentable over the

article "Topic Detection and Tracking Pilot Study Final Report" by Allan et al in

view of the article "Relevance Models for Topic Detection and Tracking" by

Lavrenko et al in view of US Patent No 7,085,755 to Bluhm et al as applied to

claim 1 above, and further in view of US PGPub 2006/0062451 to Li et al (hereafter

Li).

Referring to claim 2, while FinalReport/RelevanceModel/Bluhm teaches the

method of claim 1, wherein the inter-story similarity metric is adjusted based on at least

one of subtraction (Relevance Model: see section 4.3: Relevance Model Performance,

lines 1-13 – the symmetric clarity-adjusted KL is the similarity metric utilized; and

section 4.2: Value of Clarity-adjusted KL, line 12 – the equation uses subtraction) and

division (Relevance Model: see section 3.2: Measuring Topic Similarity, lines 31-37 –

clarity is considered to represent the adjustment), FinalReport/RelevanceModel/Bluhm

fails to explicitly disclose the limitation of the adjustment being dynamic. Li discloses a

learning procedure for text classification (see abstract), including the further limitation of

dynamically adjusting weights (see [0081]).

It would have been obvious to one of ordinary skill in the art to apply the dynamic

adjustment disclosed by Li to the metric of FinalReport/RelevanceModel/Bluhm. One

would have been motivated to do so in order to increase the accuracy of the classifier

by updating the classifier as more documents are trained.

6.   Claims 3-5 are rejected under 35 U.S.C. 103(a) as being unpatentable over the article "Topic Detection and Tracking Pilot Study Final Report" by Allan et al in view of the article "Relevance Models for Topic Detection and Tracking" by Lavrenko et al in view of US Patent No 7,085,755 to Bluhm et al as applied to claim 1 above, and further in view of US Patent No 6,584,220 issued to Lantrip et al (hereafter Lantrip et al).

Referring to claim 3, FinalReport/RelevanceModel/Bluhm teaches the further limitation wherein the inter-story similarity metric is at least one of a probability based inter-story similarity metric (see section 3.2: Measuring Topic Similarity, lines 1-14). However, FinalReport/RelevanceModel/Bluhm does not explicitly teach the further limitation wherein the inter-story similarity metric is also a Euclidean based inter-story similarity metric. Lantrip et al teaches a method of similarity metrics including a Euclidean based inter-story similarity metric (Lantrip et al: see column 4, lines 46-48).

It would have been obvious to one of ordinary skill at the time the invention was made to utilize Lantrip et al's method of a Euclidean based inter-story metric with FinalReport/RelevanceModel/Bluhm's method of a probability based inter-story metric. One would have been motivated to do so since a Euclidean based inter-story metric aids in solving the problem of how to train and cluster a corpus (Final Report: see section 1.2).

Referring to claim 4, the combination of (FinalReport/RelevanceModel/Bluhm and Lantrip et al (hereafter FinalReport/RelevanceModel/Bluhm/Lantrip) teaches the method of claim 3, wherein the probability based inter-story similarity metric is at least

one of a Hellinger, a Tanimoto, a KL divergence (Relevance Models: see section 3.2: Measuring Topic Similarity, lines 1-14 – Kullback-Leibler divergence) and a clarity distance based metric.

Referring to claim 5, FinalReport/RelevanceModel/Bluhm/Lantrip teaches the method of claim 3, wherein the Euclidean based similarity metric is a cosine-distance based metric (Lantrip et al: see column 4, lines 46-48).

7.   Claim 6 is rejected under 35 U.S.C. 103(a) as being unpatentable over the

article "Topic Detection and Tracking Pilot Study Final Report" by Allan et al in

view of the article "Relevance Models for Topic Detection and Tracking" by

Lavrenko et al in view of US Patent No 7,085,755 to Bluhm et al as applied to

claim 1 above, and further in view of the article "On-line New Event Detection and

Tracking" by Allan et al (hereafter New Event Detection).

Referring to claim 6, FinalReport/RelevanceModel/Bluhm teaches a method of

detecting new events. However FinalReport/RelevanceModel/Bluhm fails to explicitly

teach the further limitation wherein the inter-story metrics are determined based on a

term frequency-inverse story frequency model. New Event Detection discloses a

method similar to that of FinalReport/RelevanceModel/Bluhm. New Event Detection

discloses a method similar to that of FinalReport/RelevanceModel/Bluhm including the

further limitation. In particular, New Event Detection discloses a method similar to that

of claim 1, wherein the inter-story similarity metrics are determined based on a term

frequency-inverse story frequency model (see section 4.1: Detection Experiments, lines

13-15).

It would have been obvious to one of ordinary skill at the time the invention was

made to utilize New Event Detection's method of a term frequency-inverse story

frequency with FinalReport/RelevanceModel/Bluhm's method of an inter-story similarity

metrics. One would have been motivated to do so since all three articles discuss the

TDT initiative carried out by the Center for Intelligent Information Retrieval (Final Report:

see abstract; Relevance Model: see abstract; New Event Detection: see abstract).

8.   Claims 35-36 are rejected under 35 U.S.C. 102(a) as being unpatentable

over the article "Relevance Models for Topic Detection and Tracking" by

Lavrenko et al (hereafter Relevance Model) in view of US Patent No 7,085,755 to

Bluhm et al.

Referring to claim 35, Relevance Model teaches a computer-implemented

method of detecting new events comprising the steps of:

determining a first story [story S1] associated with at least one event (Relevance

Model: see section 3.2: Measuring Topic Similarity);

determining a second story associated with at least one event (Relevance Model:

see section 3.2: Measuring Topic Similarity – story S2);

determining a story-pair based on the first story and the second (Relevance

Model: see section 3.2: Measuring Topic Similarity, lines 1-14 – S1 and S2);

outputting an indicator of inter-story similarity between the first and second story

based on at least one of:

an event frequency model, story segmentation (Relevance Model: see

section 1: Introduction, lines 3-7) and a source-identified inter-story similarity

metric, wherein the event frequency model is periodically automatically updated

[the limitation recites at least one of an event frequency model, story

segmentation and a source-identified inter-story similarity metric; in this instance

an event frequency model is not being utilized].

While Relevance Model discloses a first story and a second story, Relevance Model fails to explicitly disclose the further limitations of a source-identified first story and source-identified second story. Bluhm discloses managing a large corpus of documents (see abstract), including the further limitations of a source-identified story [source/author/publisher] (see column 6, lines 33-47; column 7, lines 27-42; and column 22, lines 1-10).

It would have been obvious to one of ordinary skill in the art at the time of the invention to utilize the identification of the source of the story as disclosed by Bluhm with the stories of Relevance Model. One would have been motivated to do so in order to increase the accuracy of the results by decreasing the size of the document set searched by limiting the search to a particular source.

Referring to claim 36, the combination of Relevance Model and Bluhm discloses the method of claim 35, wherein story segmentation is based on at least one of: topic (Relevance Model: see section 1: Introduction, lines 3-7), an adjacent window and an overlapping window.

Response to Arguments

9.   Applicant's arguments with respect to the prior art rejections have been considered but are moot in view of the new ground(s) of rejection.

Contact Information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Kimberly Lovel whose telephone number is (571) 272-2750. The examiner can normally be reached on 8:00 - 4:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Cottingham can be reached on (571) 272-7079. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Kimberly Lovel
Examiner
Art Unit 2167

20 December 2007
kml

JOHN COTTINGHAM
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100